

**APPARATUS, METHOD, AND COMPUTER PROGRAM FOR PROVIDING  
MULTIMEDIA PROMPTING IN A COMMUNICATION SYSTEM**

5                   **CROSS REFERENCE TO RELATED APPLICATIONS**

[0001]       This application is related to U.S. Patent  
Application Serial No. 10/610,511 entitled "DISTRIBUTED  
CALL SERVER SUPPORTING COMMUNICATION SESSIONS IN A  
COMMUNICATION SYSTEM AND METHOD," filed on June 30, 2003,  
10   and which is incorporated herein by reference.

**TECHNICAL FIELD**

[0002]       This disclosure relates generally to  
15   communication systems and more specifically to an  
apparatus, method, and computer program for providing  
multimedia prompting in a communication system.

**BACKGROUND**

[0003] Communication systems that transport audio and video information are becoming more popular in the United States and around the world. In a conventional communication system, one participant communicates audio  
5 signals to other participants (often through a multipoint communication server or other unit) and receives audio signals from the other participants (indirectly through the server). The participants also exchange video images, which allow the participants to see one another.

10

**SUMMARY**

[0004] This disclosure provides an apparatus, method, and computer program for providing multimedia prompting in a communication system.

[0005] In one aspect, a multimedia prompting apparatus,  
5 method, and computer program generate and provide a dynamic multimedia prompt to a video client. The dynamic multimedia prompt includes at least two video clips. The multimedia prompting apparatus, method, and computer program generate the dynamic multimedia prompt by selecting  
10 at least one of the video clips in response to at least a portion of information received from the video client.

[0006] In a particular aspect, the dynamic multimedia prompt includes a first video clip that requests the information from a user of the video client, a second video  
15 clip that displays the information received from the video client, and a third video clip that requests confirmation of the information received from the video client.

[0007] Other technical features may be readily apparent to one skilled in the art from the following figures, descriptions, and claims.  
20

**BRIEF DESCRIPTION OF THE DRAWINGS**

[0008] For a more complete understanding of this disclosure and its advantages, reference is now made to the following description, taken in conjunction with the  
5 accompanying drawings, in which:

[0009] FIGURE 1 illustrates an example communication system according to one embodiment of this disclosure;

[0010] FIGURE 2 illustrates an example media application server according to one embodiment of this disclosure; and

10 [0011] FIGURE 3 illustrates an example method for providing multimedia prompting in a communication system according to one embodiment of this disclosure.

**DETAILED DESCRIPTION**

[0012] FIGURE 1 illustrates an example communication system 100 according to one embodiment of this disclosure. The system 100 shown in FIGURE 1 is for illustration only.  
5 Other embodiments of the system 100 may be used without departing from the scope of this disclosure.

[0013] A problem with conventional communication systems is that they typically lack an effective mechanism for prompting users. Moreover, conventional communication  
10 systems lack effective means for confirming whether information submitted by a user has been properly received and interpreted by the communication systems.

[0014] In the illustrated example, the system 100 includes one or more video clients 102a-102c, a network  
15 104, and a media application server ("MAS") 106.

[0015] The video clients 102a-102c represent devices used by users or subscribers during communication sessions. Communication sessions represent data conversions or conversations between at least two devices or applications  
20 over a network that involve audio and/or video information.

For example, each video client 102a-102c represents an input/output device that could include a microphone and a speaker to capture and play audio information. A video client 102a-102c could also include a camera and a display  
25 to capture and play video information.

[0016] During a communication session, the video clients 102 communicate with the MAS 106 over the network 104. As an example, the video clients 102 may transmit audio and

video information to the MAS 106 and receive audio and video information from the MAS 106. As will be appreciated, each video client 102 may be constructed or configured from any suitable hardware, software, firmware,  
5 or combination thereof for transmitting or receiving audio or video information.

[0017] The system 100 shown in FIGURE 1 illustrates various embodiments of the video clients 102. For example, the video client 102a represents a wireless video phone  
10 that communicates with the network 104 through a mobile switching center ("MSC") 108. The video client 102b represents a wired Internet Protocol ("IP") video phone that communicates directly with the network 104. The video client 102c represents a personal computer connected to a  
15 web camera 110. This represents three possible embodiments of the video clients 102. Other or additional video clients 102 may also be utilized in the system 100 of FIGURE 1. By way of illustration in FIGURE 1, each of the video clients 102a-102c is different. It will be  
20 understood, however, that the video clients 102 in the system 100 may include or represent the same or similar type of device or other combination of video clients.

[0018] The network 104 is coupled to the video clients 102, the MAS 106, and the mobile switching center 108. In  
25 this document, the term "couple" refers to any direct or indirect communication between two or more components, whether or not those components are in physical contact with each other. The network 104 facilitates communication between components of the system 100. For example, the  
30 network 104 may communicate Internet Packets ("IP"), frame

relay frames, Asynchronous Transfer Mode ("ATM") cells, Ethernet, X.25, frame relay, or other suitable information protocols between network addresses or devices. The network 104 may include one or more local area networks ("LANs"), metropolitan area networks ("MANs"), wide area networks ("WANs"), all or portions of a global network such as the Internet, or any other communication system or systems at one or more locations.

[0019] The media application server ("MAS") 106 is coupled to the network 104. The MAS 106 supports communication sessions between video clients 102 in the system 100. For example, the MAS 106 may receive requests from the video clients 102 to establish or join a communication session. The MAS 106 may also receive audio and video information from one video client 102 and communicate the information to the other video clients 102 involved in a communication session.

[0020] The MAS 106 may be constructed or configured using any hardware, software, firmware, or combination thereof for supporting communication sessions in the system 100. As an example, the MAS 106 could include one or more processors 112 that execute instructions and one or more memories 114 that store instructions and data used by the processors 112. The processor(s) 112 is generally understood to be a device that drives a general-purpose computer. It is noted, however, that other processor devices such as microcontrollers, Field Programmable Gate Arrays (FPGAs), or Application Specific Integrated Circuits (ASICs) can be used as well and achieve the benefits and advantages described herein. An example MAS 106 is shown in FIGURE 2,

which is described below and in co-pending U.S. Patent Application Serial No. 10/610,511 entitled "DISTRIBUTED CALL SERVER SUPPORTING COMMUNICATION SESSIONS IN A COMMUNICATION SYSTEM AND METHOD," filed on June 30, 2003, 5 and which is incorporated by reference herein.

[0021] Communication sessions established and managed by the MAS 106 can include communication devices other than the video clients 102a-102c. For example, a video or other client 102d is shown coupled to the network 104 through a 10 public telephone network, such as a public switched telephone network ("PSTN") 116. The video client 102d may include a conventional analog or digital telephone, a personal computer, or some other type of communication device. In embodiments where the PSTN 116 and the network 15 104 use different or incompatible protocols to communicate, a gateway 118 may be used that is coupled to the network 104 and the PSTN 116 to facilitate communication between the networks. The gateway 118 functions to translate between the different protocols used by the network 104 and 20 the PSTN 116. Although one PSTN 116 is shown in FIGURE 1 coupled to the network 104, other or additional types of public or private networks may be coupled to the network 104.

[0022] The video clients 102 and the MAS 106 could 25 support suitable standards or protocols used to set up, maintain, and terminate communication sessions between end users. As examples, the video clients 102 and the MAS 106 could communicate audio, video, or other information using the Realtime Transfer Protocol ("RTP") over User Datagram 30 Protocol ("UDP"), the International Telecommunication Union



- Telecommunications ("ITU-T") H.263 standard (video CODEC), the G.711 and G.729 audio CODEC standards, and other or additional standards or protocols. Other CODECs, such as Moving Picture Experts Group-4 ("MPEG-4"), Digital  
5 Video Express ("DIVX"), and Windows Media Video ("WMV"), can be supported by the MAS 106. In addition, signaling messages sent between the video clients 102 and the MAS 106 may include or conform with the Session Initiation Protocol ("SIP"), which is an application layer protocol for the  
10 establishment, modification, and termination of conferencing and telephony sessions over IP-based networks. As will be appreciated, other or additional protocols and configurations may be used.

[0023] In one aspect of operation, the MAS 106 provides  
15 one or more multimedia prompts to a user of a video client 102. A multimedia prompt may include a sequence of video information, which may optionally have associated audio information, that guides the user through a process. For example, a user of a video client 102 may attempt to access  
20 a particular service in the system 100, such as a movie schedule. The multimedia prompts appear on the user's video client 102 and guide the user through the various options of the movie schedule, such as locating the theatre and the desired movie.

25 [0024] In some embodiments, the MAS 106 supports both static and dynamic multimedia prompts. As an example of a static prompt, the MAS 106 could provide a multimedia prompt that requests the user of a video client 102 to enter a destination number when the user wishes to initiate  
30 a communication session. The destination number identifies

the video client 102 that the user wishes to contact. The multimedia prompt then waits for the user to enter the destination number. The user could enter the number in a variety of ways, such as by speaking the destination number  
5 or entering the destination number using a keypad on the video client 102. The multimedia prompt then requests the user to enter the user's personal identification number ("PIN"). The multimedia prompt then waits for the user to enter the PIN. When the PIN is provided, the multimedia  
10 prompt ends, and the MAS 106 initiates the communication session.

[0025] The MAS 106 could also generate and provide dynamic multimedia prompts to a video client 102. Using the same example from above, the MAS 106 could provide a  
15 multimedia prompt of a person (or other animate or inanimate object) requesting the user to enter a destination number. As the user provides each numeral of the destination number, the MAS 106 changes the multimedia prompt to physically display the numeral provided. In  
20 other words, when the user provides a numeral of the destination number, the MAS 106 changes the multimedia prompt so that it shows or displays the numeral provided by the user. When the user provides another numeral of the destination number, the MAS 106 again changes the  
25 multimedia prompt to display the second numeral on the video client 102. This continues until the entire destination number has been entered, at which point the MAS 106 changes the multimedia prompt to show or display a person requesting whether the destination number displayed  
30 by the MAS 106 is correct. If correct, the MAS 106 changes

the multimedia prompt so that it shows the person requesting the user's PIN, and the process may be repeated with additional numerals entered by the user.

[0026] This represents one possible multimedia prompt  
5 that could be generated. Another multimedia prompt could show a person requesting a destination number and then waiting for the entire number to be provided. Once the entire number is provided, the multimedia prompt could request confirmation, display the sequence of numerals  
10 detected by the MAS 106, and show the person waiting for the user to confirm that the displayed numbers were correct. A similar process could be used to obtain the user's PIN.

[0027] In addition to changing the multimedia prompt to  
15 display the various numerals of the destination number and the user's PIN, the MAS 106 could optionally provide audio information so the user hears the numerals detected by the MAS 106. In this way, the MAS 106 mixes audio, video, and text to create a dynamic multimedia prompt for a user.

[0028] In these examples, a dynamic multimedia prompt  
20 provided by the MAS 106 includes a sequence of video clips. Each video clip represents a sequence of one or more video images to be presented to and displayed by a video client 102. The video clips represent, for example, images of the  
25 person requesting the destination number, the person waiting for the destination number, each possible numeral (such as zero through nine) in the destination number, and the person requesting confirmation. The MAS 106 generates a multimedia prompt by switching between various video

clips as needed. For example, the video clips of the person requesting information or waiting may be mixed with video clips of numerals provided by the user. This allows the MAS 106 to generate any number of possible multimedia prompts and dynamically alter the prompts based on user input. While these examples describe the use of video clips to display numbers provided by the user, other video clips may include days of the week, months, years, monetary values, or any other suitable information.

10 [0029] In some embodiments, the various video clips used to generate a multimedia prompt may be encoded and stored in the MAS 106 or in a location accessible to the MAS 106 (not shown). In particular embodiments, each video clip is encoded using multiple CODECs. When a particular video client 102 accesses the MAS 106, the MAS 106 negotiates with the video client 102 and attempts to identify CODECs to be used when sending audio and video information to the video client 102. During the negotiation, the MAS 106 attempts to identify the CODEC(s) used by the video client 15 102 and the CODEC(s) used to encode the video clips. If possible, a CODEC used by the video client 102 and to encode the video clips is selected during the negotiation process. In this way, the video clips can be provided to the video client 102 without needing to decode and recode the video clips. This helps to reduce the processing load placed on the MAS 106.

[0030] Although FIGURE 1 illustrates one example of a communication system 100, various changes may be made to FIGURE 1. For example, varying video clients 102, networks 104, and servers 106 could be used in the system 100.

Also, the functionality of MAS 106, described above as being implemented on a server, could be implemented on any other computing device, such as a desktop computer or a laptop computer. In addition, FIGURE 1 illustrates one  
5 operational environment in which the various features of the MAS 106 may be used. These features could be implemented in any other suitable operating environment.

[0031] FIGURE 2 illustrates an example media application server 106 according to one embodiment of this disclosure.  
10 The MAS 106 illustrated in FIGURE 2 is for illustration only. Other embodiments of the MAS 106 could be used without departing from the scope of this disclosure. Also, while FIGURE 2 illustrates the MAS 106 operating in the system 100 of FIGURE 1, the MAS 106 may operate in other  
15 suitable systems.

[0032] In the illustrated example, the MAS 106 includes a media conductor 202, a media controller 204, two media processors ("MPs") 206a-206b, and a content store 208.

[0033] The media conductor 202 processes signaling  
20 messages received by the MAS 106. In some embodiments, the video clients 102 communicate the signaling messages directly (or via a gateway, which serves as an entrance/exit into a communications network) to the MAS 106. In other embodiments, the video clients 102  
25 communicate signaling messages indirectly to the MAS 106, such as when a Session Initiation Protocol ("SIP") application server 210 (that received a request from a video client 102) sends the signaling messages to the media conductor 202 on behalf of the video client 102. The video

clients 102 may communicate directly with the SIP application server 210 or indirectly through a gateway, such as gateway 118. The conductor 202 processes the signaling messages and communicates the processed messages  
5 to the media controller 204. As particular examples, the conductor 202 may implement SIP call control, parameter encoding, and media event package functionality.

[0034] The media controller 204 manages the operation of the MAS 106 to provide services to the video clients 102.  
10 For example, the media controller 204 may receive processed SIP requests from the conductor 202, where the requests involve conference or other calls. The controller 204 may then select the media processor 206 to handle each of the calls, support audio/video capability negotiations, enforce  
15 licenses controlling how the MAS 106 can be used, and control negotiations based on the licenses. The negotiations could include identifying the CODEC or CODECs to be used to encode and decode audio or video information during a call.

20 [0035] The media processors 206a-206b handle the exchange of audio or video information between video clients 102 involved in a call. For example, a media processor 206 could receive audio and video information from one video client 102 involved in a call, process the  
25 information as needed, and forward the information to at least one other video client 102 involved in the call. The audio and video information may be received through one or more ports 212, which couple the media processors 206 to the network 104. The ports 212 may represent any suitable  
30 structure operable to facilitate communication between the

MAS 106 and the network 104. In some embodiments, each of the media processors 206 represents a software application for specific media processing, such as interactive voice response ("IVR") media or conference media, which is  
5 executed on the MAS 106 hardware platform via the operating system.

[0036] In this example embodiment, each media processor 206 provides different functionality in the MAS 106. For example, in some embodiments, the media processor 206a  
10 provides IVR functionality in the MAS 106. As particular examples, the media processor 206a supports a voice mail function that can record and play messages or an auto-attendant function that provides a menu and directs callers to particular destinations based on their selections. The  
15 media processor 206a also supports the generation of multimedia prompts for a video client 102. The media processor 206b provides call conferencing functionality in the MAS 106, such as by facilitating the exchange of audio and video information between three or more video clients  
20 102.

[0037] The content store 208 provides access to content used by the various components of the system 100. For example, in some embodiments, the content store 208 provides access to stored voice mail messages and access  
25 codes used to initiate or join conference calls. The content store 208 also provides access to any other or additional information. In other embodiments, the content store 208 is replaced by a conventional database or other data storage facility.

[0038] A Java 2 Enterprise Edition ("J2EE") platform 214 is coupled to the MAS 106. The J2EE platform 214 allows the MAS 106 to retrieve information used to provide subscriber services in the system 100. For example, the  
5 J2EE platform 214 may provide audio announcements and video clips used by the IVR media processor 206a to produce multimedia prompts. The J2EE platform 214 represents one possible apparatus used to provide audio or other information to the MAS 106. Any other or additional device  
10 or apparatus may be used to provide the information to the MAS 106.

[0039] In a particular embodiment, various components of the MAS 106 represent software processes executed by the processor 112 of the MAS 106. While the components 202-208  
15 have been described as being executed by a MAS 106, the software processes could be executed by other computing devices such as a desktop computer. In other embodiments, the various components of the MAS 106 may be implemented in other ways, such as in hardware.

[0040] In one aspect of operation, the MAS 106 generates and provides multimedia prompts to the users of video clients 102. In some embodiments, the MAS 106 provides static or dynamic multimedia prompts to the users. The MAS  
20 106 may use any suitable mechanism to generate the dynamic prompts.  
25

[0041] As described above, one technique involves the MAS 106 combining multiple video clips together to generate a multimedia prompt that varies based on information provided by a user. In some embodiments, the IVR media



processor 206a is responsible for generating the multimedia prompts. For example, the IVR media processor 206a retrieves the needed video clips from the J2EE platform 214 or other source and provides the proper video clip to a video client 102. The IVR media processor 206a may also retrieve audio information associated with the video clip and provide the audio information to the video client 102.

The IVR media processor 206a may further synchronize the audio and video information for presentation to a user of the video client 102. In addition, the IVR media processor 206a may receive information provided by the user of the video client 102 and alter the multimedia prompt by providing particular video clips based on input from the user.

[0042] As described above, the user of a video client 102 provides information to the MAS 106 in any suitable manner, including speaking the information and using a keypad or other input device at the video client 102. If the user of a video client 102 provides information to the MAS 106 by speaking (such as when the user provides a destination number or PIN by speaking the numerals), the IVR media processor 206a performs speech recognition to identify what the user is saying.

[0043] Also, in some embodiments, the video clips and associated audio information are encoded using multiple CODECs and stored. The media controller 204 negotiates with the video client 102 during registration and attempts to identify the CODECs to be used when communicating with the video client 102. The user of a video client 102 may have preferred or desired CODECs that the user wishes to

use if possible. During the negotiation process, the media controller 204 attempts to first verify whether the preferred CODECs were also used to encode the video clips and audio information. If so, the media controller 204 may  
5 allow the video client 102 to use those preferred CODECs. If not, the media controller 204 attempts to identify other CODECs that the video client 102 supports and that were used to encode the video and audio information.

[0044] Based on the audio and video CODECs selected  
10 during the negotiation process, the IVR media processor 206a retrieves the appropriate video clips and audio information when generating a multimedia prompt. For example, if the same CODECs used by the video client 102 were also used to encode a needed video clip and associated  
15 audio information, the IVR media processor 206a retrieves and transmits the proper video clip and audio information to the video client 102 without the need to decode the information first. If the video client 102 supports CODECs that were not used to encode the video clips and audio  
20 information, the MAS 106 would function to transcode (or translate) information from one CODEC format to another CODEC format. In this way, video clients 102 can use any of a wide variety of CODECs to communicate with and receive service from the MAS 106.

25 [0045] Although FIGURE 2 illustrates one example of a media application server 106, various changes may be made to FIGURE 2. For example, any number of media processors 206 could be used in the MAS 106. Also, the functional divisions shown in FIGURE 2 are for illustration only.  
30 Various components can be combined or omitted or additional

components can be added according to particular functional designations or needs. In addition, while the components 202-208 have been described as being executed by a server, the components 202-208 may be executed by other hardware  
5 platforms, such as a desktop computer or a laptop computer.

[0046] FIGURE 3 illustrates an example method 300 for providing multimedia prompting in a communication system according to one embodiment of this disclosure. For ease of explanation, the method 300 is described with respect to  
10 the MAS 106 of FIGURE 2 operating in the system 100 of FIGURE 1. The method 300 may also be used by other suitable devices or in any other suitable system.

[0047] The MAS 106 detects a request for a service from a video client 102 at step 302. This may include, for  
15 example, the media conductor 202 receiving signaling messages from the video client 102 and the media controller 204 identifying a requested service associated with the signaling messages.

[0048] The MAS 106 identifies the CODECs to be used by  
20 the video client 102 and the MAS 106 to communicate at step 304. This may include, for example, several steps such as the media controller 204 negotiating with the video client 102 to identify the CODECs to be used. The CODECs can be identified in any suitable manner, such as by determining  
25 whether the CODECs that might be used include CODECs previously used to encode video clips and audio information and/or determining whether CODECs preferred by the video client 102 can be used to communicate with the video client 102.

[0049] The MAS 106 retrieves one or more introductory video clips and associated audio information (optional) for the requested service at step 306. This may include, for example, the IVR media processor 206a retrieving a video clip describing the service and informing the user of the video client 102 that the multimedia prompt will guide the user through the service. If needed, the IVR media processor 206a transcodes the information (video clip or audio information) before providing it to the video client 102 if the video client 102 does not support the CODECs used to encode the audio or video information. The video clips and audio information may be retrieved from any suitable location, such as from the J2EE platform 214. The MAS 106 provides the introductory video clips and audio information to the video client 102 at step 308.

[0050] The MAS 106 determines whether information is needed from the user at step 310, such as whether the service requested by the user requires information or other input from the user. If not, method 300 ends, and the MAS 106 may provide the service to the user.

[0051] Otherwise, the MAS 106 retrieves one or more video clips and associated audio information (optional) that request needed information and wait for the information at step 312. This may include, for example, the IVR media processor 206a retrieving a video clip of a person requesting the needed information and a video clip of the person waiting. The MAS 106 provides these video clips and audio information to the video client 102 at step 314.

[0052] The MAS 106 receives the requested information from the user at step 316. This may be accomplished, for example, by the user speaking the information and the IVR media processor 206a performing speech recognition to  
5 process the spoken information, the user using a keypad or other input device at the video client 102 to provide the information, or in any other suitable manner.

[0053] The MAS 106 retrieves one or more video clips and audio information (optional) confirming the requested  
10 information at step 318. This may include, for example, the IVR media processor 206a retrieving a video clip of each numeral, audio information of a person or other entity saying each numeral, or other information provided by the user. The MAS 106 provides the video clips and audio  
15 information (optional) to the video client 102 and requests confirmation that the information is correct at step 320. This may include, for example, the IVR media processor 206a providing a video clip to the video client 102 after the user says each numeral or providing a stream of video clips  
20 after the user says all numerals.

[0054] The MAS 106 determines whether additional information is needed from the user at step 322. If not, the method 300 ends, and the MAS 106 provides the requested service. Otherwise, the MAS 106 returns to step 312 and  
25 attempts to receive the needed information from the user.

[0055] Although FIGURE 3 illustrates one example of a method 300 for providing multimedia prompting in a communication system, various changes may be made to FIGURE 3. For example, FIGURE 3 illustrates the MAS 106

generating a particular dynamic multimedia prompt for a video client 102. The multimedia prompt described in FIGURE 3 includes an introductory video clip, a video clip asking for information, a video clip of a person waiting  
5 for information, video clips of user input, and a video clip of a person asking for confirmation. Other multimedia prompts that include other video clips could be generated by the MAS 106.

[0056] It may be advantageous to set forth definitions  
10 of certain words and phrases that have been used within this patent document. The terms "include" and "comprise," as well as derivatives thereof, mean inclusion without limitation. The term "or" is inclusive, meaning and/or. The phrases "associated with" and "associated therewith,"  
15 as well as derivatives thereof, may mean to include, be included within, interconnect with, contain, be contained within, connect to or with, couple to or with, be communicable with, cooperate with, interleave, juxtapose, be proximate to, be bound to or with, have, have a property  
20 of, or the like.

[0057] While this disclosure has described certain embodiments and generally associated methods, alterations and permutations of these embodiments and methods will be apparent to those skilled in the art. Accordingly, the  
25 above description of example embodiments does not define or constrain this disclosure. Other changes, substitutions, and alterations are also possible without departing from the spirit and scope of this disclosure, as defined by the following claims.